

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-249447

(43) 公開日 平成8年(1996)9月27日

(51) Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 T 1/00			G 0 6 F 15/62	3 8 0
	7/00		G 1 0 L 3/02	A
// G 1 0 L 3/02		9061-5H	G 0 6 F 15/70	4 6 0 Z
H 0 4 N 7/24			H 0 4 N 7/13	Z

審査請求 有 請求項の数 2 O L (全 7 頁)

(21) 出願番号 特願平7-52737

(22) 出願日 平成7年(1995)3月13日

(71) 出願人 000127695

株式会社エイ・ティ・アール通信システム
研究所京都府相楽郡精華町大字乾谷小字三平谷5
番地

(72) 発明者 坂口 竜己

京都府相楽郡精華町大字乾谷小字三平谷5
番地 株式会社エイ・ティ・アール通信シ
ステム研究所内

(72) 発明者 大谷 淳

京都府相楽郡精華町大字乾谷小字三平谷5
番地 株式会社エイ・ティ・アール通信シ
ステム研究所内

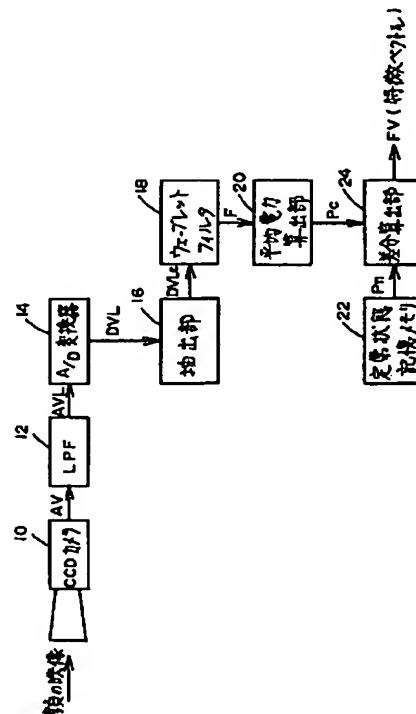
(74) 代理人 弁理士 深見 久郎 (外2名)

(54) 【発明の名称】 表情検出装置

(57) 【要約】

【目的】 顔の表情を高いロバスト性をもって実時間で正確に検出する。

【構成】 顔の映像信号DVLの中から表情が出やすい部分のみを抽出する抽出部16と、その抽出された映像信号DVLをウェーブレット変換することにより空間周波数Fを生成するウェーブレットフィルタ18と、その空間周波数Fの平均電力Pcを所定の各帯域ごとに算出する平均電力算出部20と、その算出された平均電力Pcと定常状態記憶メモリ22に予め格納されている無表情の平均電力Pnとの差分を算出することにより特徴ベクトルFVを得る差分算出部24とを備える。



【特許請求の範囲】

【請求項1】 人物の顔を撮影して映像信号を生成する撮影手段と、

前記映像信号をウェーブレット変換することにより所定の各帯域ごとに空間周波数領域の周波数信号を生成するウェーブレット変換手段と、

前記各帯域ごとに前記周波数信号の平均電力を算出する平均電力算出手段と、

前記平均電力算出手段から順次与えられる平均電力と、前記人物の顔が無表情のときにその顔から得られる対応の平均電力との差分を算出する差分算出手段とを備えた表情検出装置。

【請求項2】 前記撮影手段から与えられる映像信号の中から前記人物の顔の表情が出やすい所定領域に対応する部分を抽出して前記ウェーブレット変換手段に与える抽出手段をさらに備えたことを特徴とする請求項1に記載の表情検出装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は表情検出装置に関し、さらに詳しくは、離れた場所を結ぶコンピュータグラフィックスを利用したテレビ会議の参加者の表情検出や、実時間でコンピュータグラフィックスにより作成した顔の表情検出など、人間の顔の表情を実時間で検出する表情検出装置に関する。

【0002】

【従来の技術および発明が解決しようとする課題】 人物の顔の表情認識は、高度な符号化通信やコンピュータの能動的なインタフェースの研究において重要な基礎技術である。たとえば「大谷、北村、竹村、岸野：“臨場感通信会議における3次元顔画像の実時間表示”，信学技報，HC-92-61，pp.23-28(1993.1.)」に開示されているようなコンピュータ制御による通信会議システムにおいては、表情認識が実現されれば、伝送すべき情報の高圧縮率が達成できるとともに、より高度な強調作業などへの応用が期待できる。

【0003】 上述した臨場感通信会議における従来の表情検出方法では、顔の数カ所にマーカーを貼り付け、そのマーカーの動きを追跡することによって顔の表情を検出していた。しかし、このような方法ではマーカーなどの検出補助用具を必要とするため、表情検出を実行するための前準備に時間がかかるという問題があった。また、ロバスト（堅牢）性が低いため、ヘルメットのずれや脈拍などの外乱要因によって誤動作を生じやすいという問題もあった。

【0004】 また、コンピュータビジョンの分野では、顔の特徴部位（目、鼻、口など）の位置や形状を画像から抽出し、その測定量に基づいて顔の表情認識を試みている例が多い。しかし、画像の特徴量として各部位の形状や位置を写真などの静止画像から正確に抽出すること

は、撮影条件や被験者の面相に強く依存するため、非常に困難であると同時に、複雑な処理アルゴリズムが必要となり、実時間処理を実現するための障害となっていた。

【0005】 ところで、時間変化に着目した顔の表情認識についてはいくつかの報告がなされている。たとえば「M. Rosenblum, Y. Yacoob and L. Davis: “Human Emotion Recognition from Motion Using a Radial Basis Function Network Architecture”, Proceedings of the IEEE Workshop on Motion of Non-rigid and Articulated Objects, pp.43-49(1994)」では、動画像中の表情による顔表面の変化に着目し、オプティカルフローを用いた認識を試みているが、非常に多次元の特徴となるため、隠れマルコフモデル（HMM: Hidden Markov Model）による認識には向かず、処理が複雑で計算量が多いという問題があった。

【0006】 HMMを利用した動画像認識の例として、たとえば「大和、大谷、石井：“隠れマルコフモデルを用いた動画像からの人物の行動認識”，信学会論文誌D-II, Vol. J76-D-II, No. 12, pp2556-2563(1993.12.)」では、テニスプレーヤーの動き認識を試みているが、画像の特徴化に用いている輝度値のメッシュ特徴が人物表情のように類似度が高い場合は必ずしも有効でなく、類似したカテゴリへの分類には弱いなどの課題が残っていた。

【0007】 この発明は上記のような問題点を解決するためになされたもので、ロバスト性の高い表情検出装置を提供することを目的とする。

【0008】 この発明の他の目的は、顔にマーカーを貼り付けるなどの前準備を行なうことなく、実時間で顔の表情を検出できる表情検出装置を提供することである。

【0009】 この発明のさらに他の目的は、処理が単純で計算量が少ない表情検出装置を提供することである。

【0010】 この発明のさらに他の目的は、HMMを適用することにより高い表情認識率を得ることが可能な表情検出装置を提供することである。

【0011】

【課題を解決するための手段】 請求項1に係る表情検出装置は、撮影手段、ウェーブレット変換手段、平均電力算出手段、および差分算出手段を備える。撮影手段は、人物の顔を撮影して映像信号を生成する。ウェーブレット変換手段は、映像信号をウェーブレット変換することにより所定の各帯域ごとに空間周波数領域の周波数信号を生成する。平均電力算出手段は、各帯域ごとに周波数信号の平均電力を算出する。差分算出手段は、平均電力算出手段から順次与えられる平均電力と、人物の顔が無表情のときにその顔から得られる対応の平均電力との差分を算出する。

【0012】 請求項2に係る表情検出装置は、上記請求項1の構成に加えて、抽出手段をさらに備える。抽出手段は、映像手段から与えられる映像信号の中から人物の

3

顔の表情が出やすい所定領域に対応する部分を抽出してウェーブレット変換手段に与える。

【0013】

【作用】請求項1に係る表情検出装置においては、人物の顔が撮影されて映像信号が生成される。その生成された映像信号はウェーブレット変換されることにより所定の各帯域ごとに空間周波数領域の周波数信号が生成され、さらに各帯域ごとにその生成された周波数信号の平均電力が算出される。そして、その算出された平均電力と、無表情の顔から得られる対応の平均電力との差分が算出される。このように映像信号を一旦ウェーブレット変換しているため、ロバスト性の高い表情検出が可能となる。

【0014】請求項2に係る表情検出装置においては、上記請求項1の作用に加えて、表情が出やすい部分だけを映像信号の中から抽出してウェーブレット変換しているため、より正確な表情検出が可能となる。

【0015】

【実施例】以下、この発明に係る表情検出装置の一実施例を図面を参照して詳しく説明する。なお、図中同一符号は同一または相当部分を示す。

【0016】図1は、この発明の一実施例による表情検出装置の全体構成を示すブロック図である。図1を参照して、この表情検出装置は、人間の顔を撮影するためのCCD (Charge Coupled Device) カメラ10と、CCDカメラ10から与えられるアナログ映像信号AVの中から高域成分を除去するローパスフィルタ(LPF) 12と、ローパスフィルタ12から与えられるアナログ低域映像信号AVLをデジタル低域映像信号DVLに変換するA/D変換器14と、A/D変換器14から与えられるデジタル低域映像信号DVLの中から後述する予め定められた部分だけを抽出する抽出部16と、抽出部16によって抽出されたデジタル低域映像信号DVLをウェーブレット(Wavelet)変換することにより所定の各帯域ごとに空間周波数領域の周波数信号Fを生成するウェーブレットフィルタ18と、各帯域ごとにその生成された周波数信号Fの平均電力Pcを算出する平均電力算出部20と、この平均電力Pcに対応して無表情のときの平均電力Pnが予め格納されている定常状態記憶メモリ22と、それら平均電力PcおよびPnの差分を算出することにより特徴ベクトルFVを生成する差分算出部24とを備える。

【0017】次に、この表情検出装置の動作について説明する。まず被験者の顔の正面にCCDカメラ10を取付ける。このCCDカメラ10は、常に定位置から被験者の顔を撮影するように被験者に装着されたヘルメットなどに固定される。図2は、CCDカメラ10によって撮影した被験者の顔の画像を示す。

【0018】表情を検出するのに先立って、予め無表情のときの平均電力Pnを定常状態記憶メモリ22に格納

4

しておく。この平均電力Pnの詳細は後述する。

【0019】表情を検出するに当たっては、まずCCDカメラ10によって被験者の顔が撮影され、それにより生成されたアナログ映像信号AVがローパスフィルタ12に与えられる。ローパスフィルタ12はその与えられたアナログ映像信号AVの中から高域成分を除去し、低域アナログ映像信号AVLをA/D変換器14に与える。一般に人間の顔には、毛穴、凹凸、産毛などがあるために、このアナログ映像信号AVにはそれらに起因する高域成分が大量に含まれているが、このような高域成分は顔の表情を検出するためには全く不要なものである。また、SN比の低いアナログ映像信号AVには大量のノイズ成分が含まれているが、このようなノイズ成分のほとんどは高い周波数を有している。したがって、ローパスフィルタ12は、表情の検出を行なうために必要な低域成分のみを通過させることによってこの装置の誤動作を防止するためのものである。

【0020】A/D変換器14は、その与えられる低域アナログ映像信号AVLを低域デジタル映像信号DVLに変換する。この低域デジタル映像信号DVLは抽出部16に与えられ、この低域デジタル映像信号DVLの一部DVL_eが抽出される。この実施例では、図2に示されるように目および眉を含む検出領域161と口を含む検出領域162とが切り出される。つまり、映像信号DVLの中から検出領域161および162に対応する部分DVL_eのみが抽出される。このように目、眉および口を検出領域と定めたのは、これらが表情による変化が顕著に現れる部位だからである。したがって検出領域は上述した領域に限定されることなく、たとえば皺の生じやすい額を検出領域に設定してもよい。

【0021】このようにして抽出された映像信号DVL_eはウェーブレットフィルタ18に与えられ、ここでウェーブレット変換が行なわれる。ウェーブレット変換とは、画像の周波数領域への変換手法の1種であり、高周波ほど変換基底幅が狭くなるオーバーラップ変換と考えられる。画像を周波数領域に変換するウェーブレットフィルタ18は、帯域分割フィルタと考えることができ、図3に示すようなフィルタバンクを多段階に組合せることによって実現することができる。図3に示されるように、1つのフィルタバンクは、水平方向分割用のローパスフィルタ26およびハイパスフィルタ28と、垂直方向分割用のローパスフィルタ34、42およびハイパスフィルタ36、44と、ダウンサンプラ30、32、38、40、46、48とを備える。

【0022】抽出部16からの映像信号DVL_eは、原イメージとしてローパスフィルタ26およびハイパスフィルタ28に与えられる。ローパスフィルタ26を通過した信号はダウンサンプラ30によって圧縮され、さらにローパスフィルタ34およびハイパスフィルタ36に与えられる。ローパスフィルタ34を通過した信号はダ

5

ウンサンブラ38によって圧縮され、さらに次段のフィルタバンクに与えられる。ハイパスフィルタ36を通過した信号はダウンサンブラ40によって圧縮され、図4に示された周波数帯域183に現れる。

【0023】他方、ハイパスフィルタ28を通過した信号はダウンサンブラ32によって圧縮され、さらにローパスフィルタ42およびハイパスフィルタ44に与えられる。ローパスフィルタ42を通過した信号はダウンサンブラ46によって圧縮され、図4に示された周波数帯域182に現れる。ハイパスフィルタ44を通過した信号はダウンサンブラ48によって圧縮され、図4に示された周波数帯域181に現れる。

【0024】次段のフィルタバンクに与えられた映像信号は同様に、図4に示された3つの周波数帯域184~186に分割される。さらにその次の段のフィルタバンクに与えられた映像信号も同様に、3つの周波数帯域187~189に分割される。この終段のフィルタバンクにおけるローパスフィルタ34を通過した信号はダウンサンブラ38によって圧縮され、図4に示された周波数帯域190に現れる。したがって、周波数帯域181には最高の空間周波数が含まれ、周波数帯域190には最低の空間周波数が含まれる。

【0025】図5は、眉および目を含む検出領域161の映像信号が上記のような3段のフィルタバンクを通過した場合の画像を示す。この画像は可視化するために各帯域ごとにバイアスをかけている。

【0026】ウェーブレットフィルタ18中のローパスフィルタ26, 34, 42およびハイパスフィルタ28, 36, 44には、完全可逆性、線形位相に近い位相特性、急峻な遮断特性、周波数応答の直交性などの条件が必要とされるが、これらの条件にはトレードオフがある。この実施例では、遮断特性は比較的低いが直線位相特性を持った32次の直交ミラーフィルタが用いられる。

【0027】図5に示されたようにウェーブレットフィルタ18を通過した画像では、その周波数特性とともに画像上での位置情報が保存されている。この情報量は原画像と同一であるので、これから特徴を得ることは困難である。また、このまま特徴として用いたとすると、CCDカメラ10の位置の変動や個人差の影響を強く受けてしまう。そこで、この発明では各帯域ごとの平均電力の増減のみを特徴とする手法が採用されている。したがって、ウェーブレットフィルタ18によって各帯域ごとに生成された空間周波数Fは平均電力算出部20に与えられ、ここで空間周波数Fの平均電力Pcが各帯域ごとに算出される。

【0028】表情が表出される際には顔の構成要素の形状や大きさ、傾き等が変化する。この変化が各々の周波数帯域内に与える影響を特徴とする。たとえば被験者が目を閉じた場合は、それまで電力の存在していた水平周

6

波数成分が減少し、逆に電力の少なかった垂直高周波成分が増加する。

【0029】表情が変化した場合、高周波数帯域181~183では顕著な電力変化は認められないが、低周波数帯域184~189では顕著な電力変化が認められる。これは、ウェーブレットフィルタ18を通過した後の高域成分には主に画像のエッジ情報が含まれ、低域成分には顔の構成要素の形状情報が含まれているからである。このことは、表情変化に伴うエッジ方向や強度の変化に比べ、形状そのものの変化のほうが特徴として有効であることを示す。したがって、この実施例では周波数帯域184~189内の平均電力Pcが好ましく用いられる。

【0030】図6は、被験者が驚いた場合に口を含む検出領域162から得られた周波数帯域185の平均電力Pcの変化を示すグラフである。この表の横軸はフレーム数を示し、縦軸は平均電力を示す。驚きの表情が現れた場合、口が縦に開かれるため、帯域185に現れる水平の低域成分の電力が顕著に増加する。

【0031】図7は、被験者が瞬きをした場合に眉および目を含む領域161から得られる周波数帯域185および186の平均電力の変化を示すグラフである。図7に示されるように、瞬きが行なわれると、目が細くなるので帯域185に現れる水平の低域成分の平均電力が瞬間的に減少し、帯域186に現れる垂直の低域成分の平均電力が瞬間的に増加する。

【0032】このように、ウェーブレット変換により得られた空間周波数の低域成分における平均電力は、顔の表情に応じて変化する。平均電力算出部20において算出されたこのような平均電力Pcは各フレームごとに差分算出部24に与えられる。差分算出部24では、この算出された平均電力Pcと定常状態記憶メモリ22から与えられる無表情の平均電力Pnとの差分が算出される。ここで、定常状態記憶メモリ22には、算出された平均電力Pcに対応するような平均電力Pnが格納されていなければならない。したがって、表情を検出するのに先立って、まず無表情の顔を撮影し、上記と同様にウェーブレット変換を行ない、さらに帯域184~189の平均電力を算出して定常状態記憶メモリ22に予め格納しておく必要がある。したがって、算出された平均電力Pcと無表情の平均電力Pnとの差分が差分算出部24から特徴ベクトルFVとして出力される。特徴ベクトルFVは、無表情を基準とした顔の変位を示す。

【0033】以上のようにこの実施例によれば、顔の映像信号を一旦ウェーブレット変換しているため、ロバスト性が高く、画面内で顔が多少移動しても、あるいは多少照明条件が異なっても、表情を正確に検出することができる。また、単純な構成のフィルタによってウェーブレット変換が行なわれるため、処理が簡潔で計算量が少ない。そのため、実時間で表情の検出を行なうこと

ができる。

【0034】また、表情の出やすい眉、目および口を含む検出領域161、162のみを抽出しているため、表情変化と無関係な顔の動きによって誤動作を引起すことはなく、より正確な表情検出が可能となる。

【0035】また、この実施例によって得られた特徴ベクトルFVをベクトル量子化し、HMMを用いて表情の認識を行なうことも可能である。このようにHMMによる時間変化と組合せれば、より高い認識率を得ることが可能となる。

【0036】

【発明の効果】請求項1に係る表情検出装置によれば、顔の映像信号をウェーブレット変換し、それにより得られた空間周波数の平均電力を無表情のときの平均電力と比較するようにしているため、ロバスト性の高い表情検出が可能である。

【0037】請求項2に係る表情検出装置によれば、表情の出やすい映像信号のみを抽出しているため、より正確な表情検出が可能である。

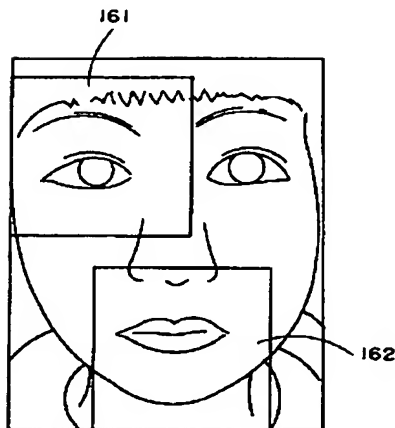
【図面の簡単な説明】

【図1】この発明の実施例による表情検出装置の全体構成を示すブロック図である。

【図2】図1中のCCDカメラによって撮影された画像であり、その中には図1中の抽出部によって抽出される検出領域が示される。

【図3】図1中のウェーブレットフィルタを構成する1

【図2】



つのフィルタバンクの構成を示すブロック図である。

【図4】図1中のウェーブレットフィルタによって得られた空間周波数の画像の構成図である。

【図5】図2に示された眉および目を含む検出領域の映像信号が実際に図1中のウェーブレットフィルタを通過した場合における空間周波数の画像である。

【図6】被験者が怒った場合に図1中の平均電力算出部から得られる平均電力の変化を示すグラフである。

10 【図7】被験者が瞬きをした場合に図1中の平均電力算出部から得られる平均電力の変化を示すグラフである。

【符号の説明】

10 CCDカメラ

12 ローパスフィルタ

14 A/D変換器

16 抽出部

18 ウェーブレットフィルタ

20 平均電力算出部

22 定常状態記憶メモリ

24 差分算出部

20 181~190 周波数帯域

AV アナログ映像信号

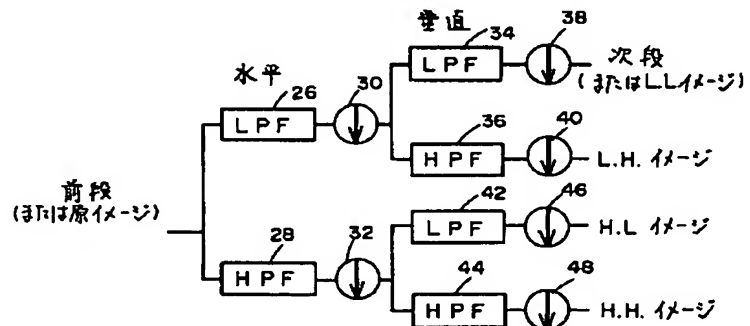
AVL 低域アナログ映像信号

DVL, DVL e 低域デジタル映像信号

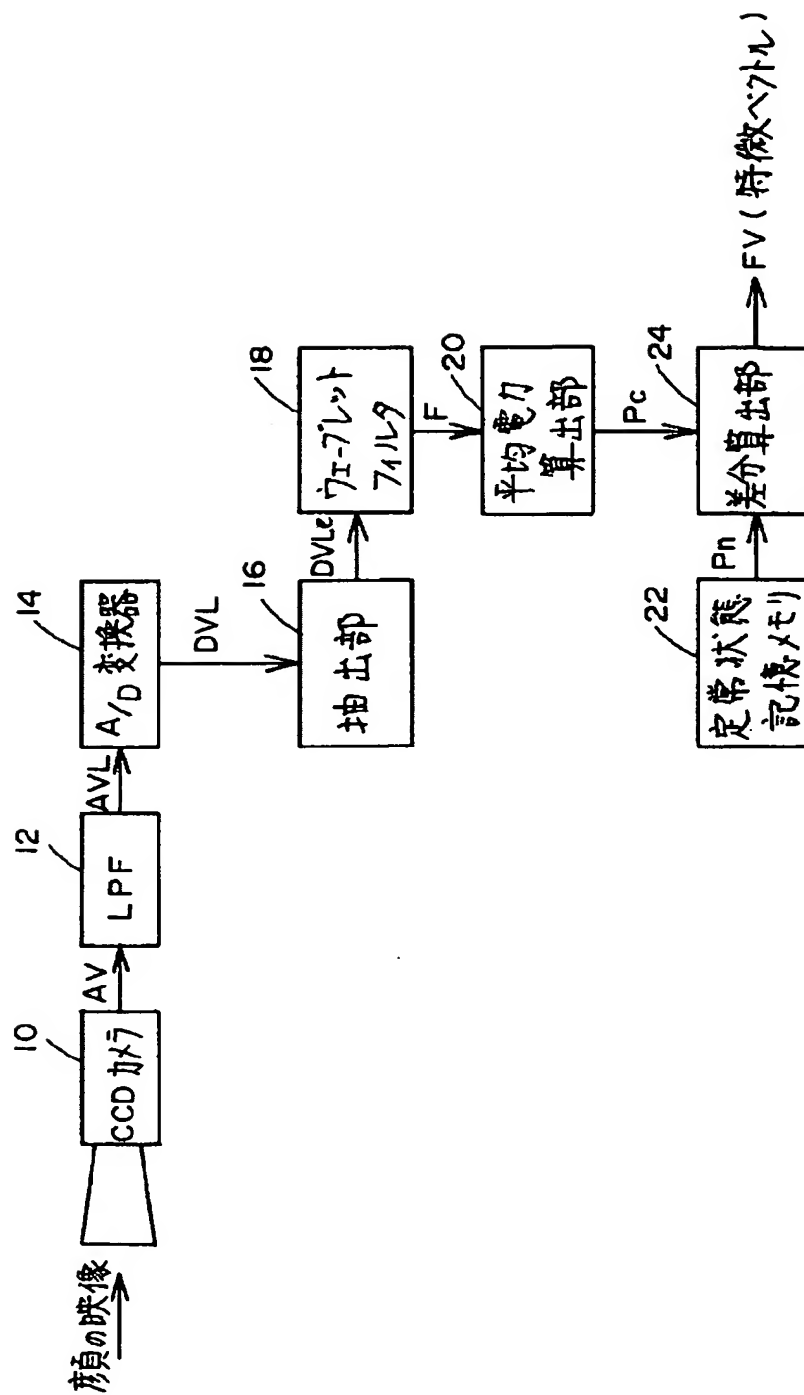
F 空間周波数

Pc, Pn 平均電力

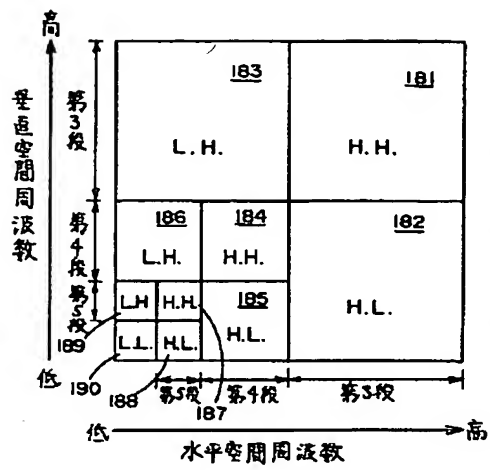
【図3】



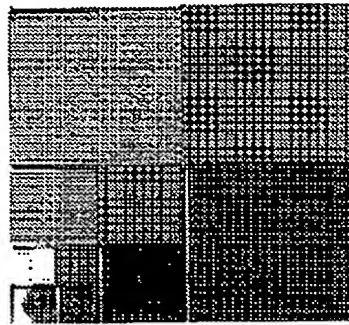
【図1】



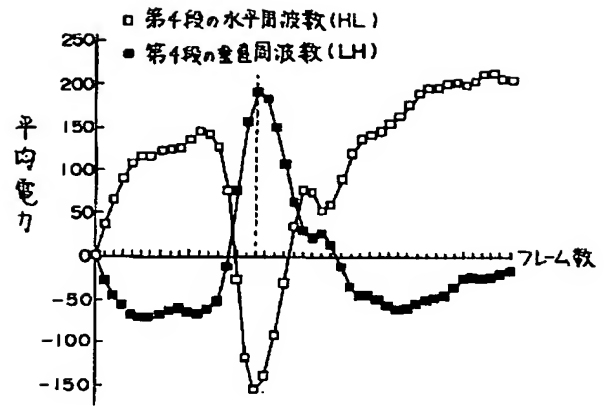
【図4】



【図5】



【図7】



【図6】

